Prevention or Identification of Web Intrusion via Human Computer Interaction Behaviour – A Proposal

Hugo Gamboa[1], Ana Fred[2] and António Alves Vieira[1]

[1]Escola Superior de Tecnologia de Setúbal
Campus do IPS Estefanilha Setúbal, Portugal
Tel +351 265790000, Fax +351 265721869
[2] Instituto de Telcomunicações
Instituto Superior Técnico IST - Torre Norte
Av. Rovisco Pais, 1 1049-001, Lisbon, PORTUGAL
Tel: +351 21 8418469 Fax: +351 21 8418472

{hgamboa,avieira}@est.ips.pt; afred@lx.it.pt

ABSTRACT

The present work proposes a new technique for the identification or prevention of intrusion in web applications via the monitoring of the user interaction behaviour. We report preliminary results in a verification task based on a user claiming his identity and being accepted or detected as an intruder after some time of user interaction monitoring. We describe the acquisition system that enables the remote monitoring of the user human computer interaction and the recognition system that detects an intrusion in the system, and present some preliminary results.

1.0   INTRODUCTION

Malicious intrusion in Internet sites either with the intention of disrupting the system or to steal information can threat the normal operation of the Web. We propose a new biometric technique capable of monitoring the user behaviour in a web site in order to prevent or detect intrusion by validating the user identity claim, normally made in a logon page. The system is based on the learnt human computer interaction behaviour of the genuine users.

We developed an acquisition system, called Web Interaction Display and Monitoring (WIDAM) [3] , that collects the user interaction data by recording the mouse movements, clicks and key presses, among other interaction events, while the user is browsing a web page.

The biometric system uses this last system to verify the identity of a, while he is navigating in the web page. The classification is based on statistical pattern recognition models, and is done after a period of interaction. This period can be selected in order to define the security level of the biometric system. In our preliminary results we obtained a level of security similar to other behavioural biometrics techniques if a period of 60 seconds of interaction is collected. This methodology introduces a possibility of applying an biometric layer to web systems with the present technology. We will now introduce some terminology used in the biometric area.

Biometric systems can be divided in two types [7]: (1) Identity verification (or authentication) occurs when a user claims who he is and the system accepts (or declines) his claim; (2) Identity identification (sometimes called search) occurs when the system establishes a subject identity (or fails to do it) without any prior claim. Biometric techniques can also be classified according to the type of characteristics explored : (1) physiological --- a physiological trait tends to be a stable physical characteristic, such as

		Form Approved
Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **25 OCT 2004**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Prevention or Identification of Web Intrusion via Human Computer Interaction Behaviour A Proposal	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Escola Superior de Tecnologia de Setúbal Campus do IPS Estefanilha Setúbal, Portugal**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM201977, Systems, Concepts and Integration Methods and Technologies for Defence against Terrorism (Systemes, concepts, methodes d'integration et technologies pour la luttre contre le terrorisme)., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**30**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

finger print, hand silhouette, blood vessel pattern in the hand, face or back of the eye. (2) behavioural --- a behavioural characteristic is a reflection of an individual's psychology.

The evaluation of a biometric technique requires the definition of metrics that can be used for the comparison of performance among different techniques [8], typically: False rejection rate (FRR) --- rate of accesses where a legitimate user is rejected by the system; False acceptance rate --- rate of accesses where an intruder is accepted by the system; Equal error rate (EER) --- the value at which FAR and FRR are equal.
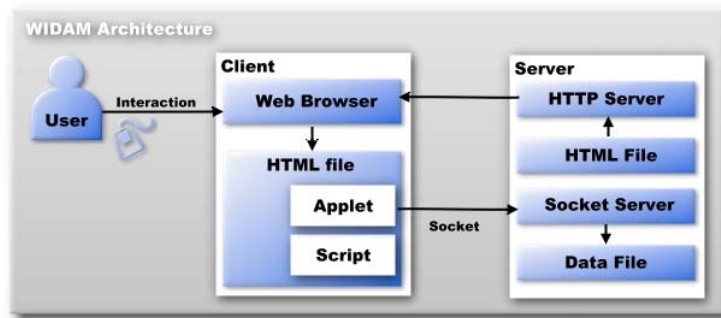

Figure 1 The WIDAM Architecture

In this paper we propose both a web based user interaction monitoring system called Web Interaction Display and Monitoring, WIDAM, and a new behavioural biometric technique based on web interaction via a pointing device, typically a mouse pointer. The normal interaction through this device is analysed for extraction of behavioural information in order to link an identification claim to an individual.

In the following section we present the user interaction acquisition system, WIDAM. In section 3 we describe the authentication system, focusing on the sequential classifier. Section 4 presents experimental results obtained using the collected data. Conclusions are presented in section 5.

2.0 THE ACQUISITION SYSTEM

The acquisition system, WIDAM, (this system is presented with more detail in [3]) enables the user interaction monitoring, analysis and display on web pages. The system can be called as a remote display system that enables the synchronous and asynchronous observation of the user interaction.

WIDAM allows the usage of an interaction recording system directly over a web page, based on the Document Object Model [5] (DOM) of the web page. The system works in a normal web browser with java and JavaScript capabilities, without the need of any software installation. WIDAM is a lightweight networked application using low bandwidth comparatively to image based remote display systems. The WIDAM Architecture is composed by a client and server applications, as depicted in figure 1. The user accesses the monitored WIDAM web page via a web browser that connects to the server. Then, the server sends back to the user a web page that is capable of monitoring and displaying the user interaction. This page creates a connection to the server and selects one of the services provided by WIDAM. Then the client and the server exchange messages using a specific protocol.

The client works in any web browser capable of executing JavaScript code and Java Applets, independent of the operating system. When the users enter into a page of the WIDAM system, an applet is launched. This applet creates a socket connection that enables the message passing from, and to the server. The client loads the html page and sends an handshaking message through the open socket, specifying which type of service is requested.

In the case of a Recording Service or Synchronous Monitoring Service, the script sends a request to the browser, asking for notification of the user interface events (a sub set of the events from the Document Object Model Events [9] listed in table 1).

ID	Event handler	Event cause
0	onMouseMove	The user moves the cursor.
1	onMouseDown	The user presses a mouse button.
2	onKeyPress	The user presses a key.
3	onUnload	The user exits a document.
4	onMove	The window is moved.
5	onSelect	The user selects some text.
6	onResize	The window is resized.
7	onBlur	The window loses focus.
8	onFocus	The window receives focus.

Table 1 DOM events captured by WIDAM

For the purpose of the intrusion detection technique being developed, the WIDAM system operated in a recording mode, over a web page with the memory game: a grid of tiles, each tile having associated a hidden pattern, which is shown for a brief period of time upon clicking on it; the purpose of the game is to identify the matching tiles. The WIDAM system presents a web page to the user, asking for his identification (name, and a personal number). Then the system presents an *interaction acquisition page* with the memory game (that could be any html web page), depicted in figure 2. This page is monitored by the WIDAM application that records all the user interaction in a file stored in the web server. Figure 3 shows a graph of a user interaction while playing an entire memory game. The graph is produced by joining every sequential mouse movement with lines and using a cross mark to indicate a mouse click.
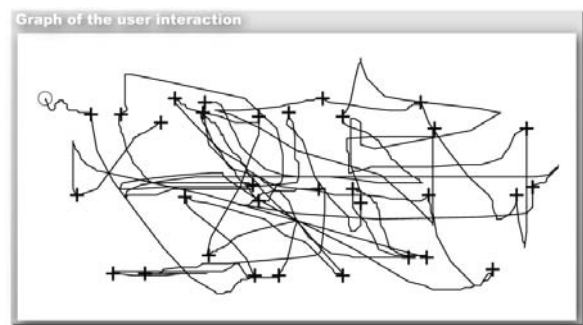

**Figure 2 Interaction test page: the
memory game.**

**Figure 3 Graph of the user interaction in
the game.**

3.0 THE AUTHENTICATION SYSTEM

An experimental system - the authentication system - was developed to verify the possibility of detecting an intrusion or invalid identity claim, using the computer interaction information, specifically based on mouse movements performed between successive clicks, which we will call a *stroke* (see figure 4).

Figure 5 presents the acquisition and recognition systems and its respective building blocks. The acquisition system was addressed in the previous section. The recognition system comprises the following modules: (a) feature extraction; (b) feature selection; (c) parametrical learning; (d) statistical sequential classifier. The recognition system reads the interaction data from the stored data files produced by the acquisition system. The interaction data passes a feature extraction procedure, creating a 63-dimensional

vector, exploring both spatial (related to angle and curvature) and temporal (related to duration, position, velocity and acceleration) characteristics of the strokes. More details can be found in [4].
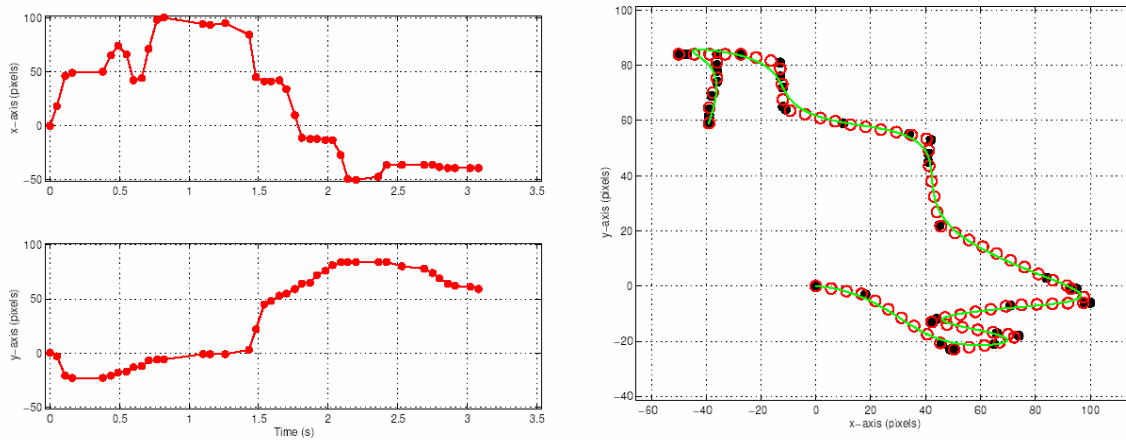

Figure 4 Left**: Input signals generated by the mouse move events.** Right**:The x-y representationof
the signals. Black circles represent the input sampled points generated by mouse move events.
White circles represent linearly (equidistant) interpolated points. The line represent the
smoothed spline interpolation.)**

The system has an enrolment phase, where the global set of extracted features are used in an algorithm that selects a set of "best" features for each user, using the equal error rate as performance measure (feature selection block in figure 5), using the Sequential Forward Selection (SFS) [6] that selects the best single feature and then adds one feature at time to a the vector of previously selected features. The algorithm stops when the equal error rate does not decrease.

The classification rule assumes a statistical model for the feature vectors. The learning phase consists of the estimation of the probability density functions, $p(X)$ (where X is the feature vector of a stroke), from each user's data. Considering that each user constitutes a recognition class, and assuming statistical independence between features, $p(X)$ factorizes into $p(X|user) = \Pi p(x_i|user)$. We use as parametrical model for $p(x_i|user)$ the *weibull* [1] distribution ($p(x|a,b) = a\, b\, x^{(b-1)}\, e^{(-a\, x^b)}$). Given the data from one user and one feature, maximum likelihood estimates of the parameters a and b are obtained.

The classifier's purpose is to decide if a user is whom he claims to be, based on the patterns of interaction with the computer. We consider that the ith user is denoted by the class w_i, $i=1, \dots, L$, and L is the number of classes. As defined before, a feature vector is associated with one stroke. Given a sequence of n_s consecutive strokes executed by the user, w_i, interaction information is summarized in the vector $X = X^1 \dots X^n$, consisting of the concatenation of the feature vectors associated with each stroke.

The classifier will decide to accept or reject the claimed identity based on two distributions: the genuine distribution $p(X|w_i)$, and the impostor distribution $p(X|w_j)\, j \neq i$ that is based on a mixture of distributions (weibull distributions), one for each other user not equal to i, weighted by L, the number of users.
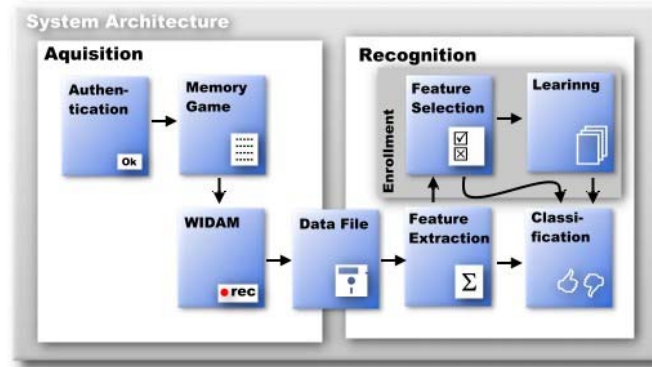
Figure 5 Authentication system architecture.

Since $p(w_i|X)$ represents an estimate of the probability of the classification being correct, we establish a *limit*, λ to select one of the decisions, using the decision rule:

$$Accept(\mathbf{X} \in w_i) = \begin{cases} true & \text{if } p(w_i|\mathbf{X}) > \lambda \\ false & \text{otherwise} \end{cases}$$

To present result about the classifier performance we adjust λ to operate in the equal error rate point, establishing the error of detection of intrusion equal to the error of accepting an intrusion.

4.0 RESULTS

We asked 50 volunteers (engineering students) to use the developed system, playing several memory games during about 10-15 minutes. This way, we created an interaction repository of approximately 10 hours of interaction, providing more than 400 strokes per user. The acquisition system monitors the pointing device with a sample rate of 50 times per second, producing messages form the client to the server that require approximately 1 Kbytes/s (950 bytes per second) as the maximum bandwidth. For instance, the ten hours of interaction occupies 36 Mbytes of disk space.

In order to use the same number of strokes per user in the tests performed, we randomly selected 180 strokes from each user. The set of strokes was divided into two equal parts, one for the training phase and other for the testing phase. Using the training set we learnt the parametrical distribution $p(x_i|user)$ for each user and each feature. Feature selection used the same data set and was tuned for each user, based on the performance of the system using sequences of 10 strokes. When testing the system for one user, we consider an intruder as one of the other users. The test function returns the equal error rate given N sequences of strokes of length l using the classifier tuned for user i. The input sequence of strokes of a test is composed of $N/2$ strokes randomly sampled from the testing set of the user, and $N/2$ strokes randomly sampled from the testing sets of all the other users.

One of the free variables of the system is the number of strokes that the system will use in the verification task. Bootstrap [2] estimates of the system performance as a function of the sequence stroke length (from 1 to 100 strokes) was obtained using 10000 bootstrap samples from the test set. The mean duration of a stroke is approximately 1 second. A graphical display of these results is shown in figure 6. As shown, the mean value and the standard deviation of the EER progressively tends to zero as more strokes are added to the decision rule. This illustrates the refinement of the performance obtained by the sequential classifier.

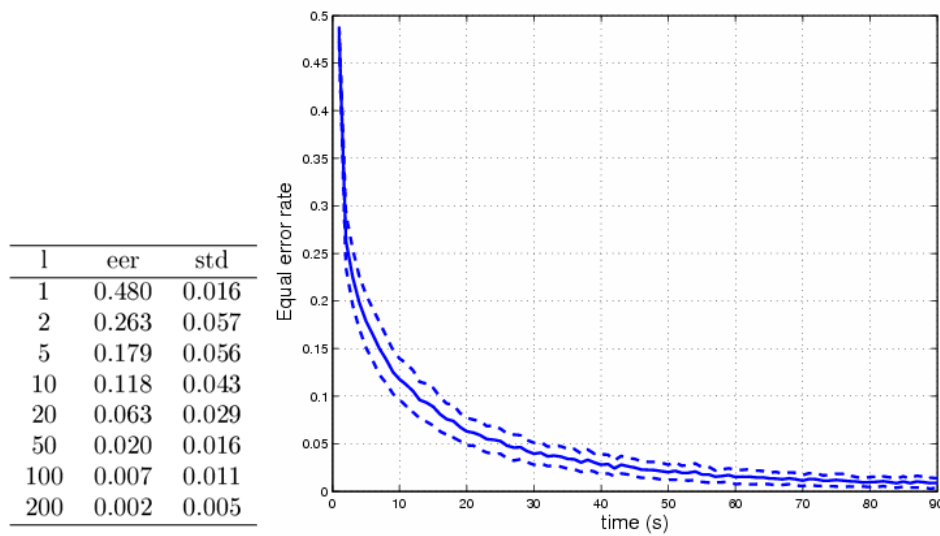| l	eer	std
1	0.480	0.016
2	0.263	0.057
5	0.179	0.056
10	0.118	0.043
20	0.063	0.029
50	0.020	0.016
100	0.007	0.011
200	0.002	0.005

**Figure 6 Equal error rate results of the verification system.   The solid line is the mean of the
equal error rate of all users. The dashed lines are the mean plus and minus half standard
deviation**

Table 2 presents EER values reported in the literature for several biometric techniques [10]. Preliminary
results show that the proposed technique, based on behavioural information extracted from the interaction
with the computer, can achieve comparable performances with other biometric techniques.

Biometric technic	Equal error rate
Retinal Scan	1:10 000 000
Iris Scan	1:131 000
Fingerprints	1:500
Hand Geometry	1:500
Signature Dynamics	1:50
Voice Dynamics	1:50
$30s$ of User Interaction	1:50
$60s$ of User Interaction	1:100
$90s$ of User Interaction	1:200

Table 2 Comparison between several biometric techniques

5.0 CONLUSION

The human computer interaction behaviour used in this work carries information related to the author
identity. We have studied this source of information in order to develop a system capable of detection or
prevention of web intrusion by the means of a behavioural biometric layer in web applications.

The acquisition of the interaction was based on the developed WIDAM system that works over the World
Wide Web collecting the mouse clicks and movements among other interaction events.

The intrusion detection is based on a new biometric technique based on statistical pattern recognition
methods. The biometric layer rejects the user claim past entering in the system by logging thru the security
page of the web application. After an established period, the claim is validated or rejected via the
recognition system that compares the acquired user interaction with the features of the previously acquired
data from the user.

The prevention of intrusion in web sites is accomplished by the immediate rejection of automatic scripts that try to gain access, commonly called bots. The bots do not present the typical human computer interaction, since the scripts interact directly with the server without any usage of interaction devices. This absence of events from a typical human can easily be used to deny access to any automatic malicious (or not) script.

The results of the tests with 50 users and a total of 10 hours of interaction showed that this technique can be applied to produce a web based behavioural biometric system to identify and prevent intrusion in web applications. The performance results are comparable to some of the behavioural biometric techniques an inexpensive technique that operates remotely using the human-computer interaction behaviour.

6.0   REFERENCES

[1]   Robert B. Abernethy. The New Weibull Handbook. Robert B. Abernethy, 2000.

[2]   Bradley Efron and Robert J. Tibshirani. An Introduction to the Bootstrap. Chapman & Hall, 1993.

[3]   Hugo Gamboa and Vasco Ferreira. WIDAM - Web Interaction Display and Monitoring. In Proceedings of the 5th International Conference on Enterprise Information Systems, volume 4, pages 21-27, 2003.

[4]   Hugo Gamboa and Ana Fred. An Identity Authentication System Based On Human Computer Interaction Behaviour. In Proceedings of the 3rd International Workshop on Pattern Recognition in Information Systems, pages 46-55, 2003.

[5]   Arnaud Le Hors, Philippe Le Hgaret, and Lauren Wood. Document object model level 2 core. Technical report, W3C, 2000.

[6]   Anil K. Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1), 2000.

[7]   Vaclav Matyas Jr and Zdenek Riha. Biometric authentication systems. Technical report, ECOM-MONITOR, 2000.

[8]   Tony Manseld and Gary Roethenbaugh. 1999 glossary of biometric terms. Technical report, Association for Biometrics, 1999.

[9]   Tom Pixley. Document object model (dom) level 2 events specification. Technical report, W3C, 2000.

[10] Thomas Ruggles. Comparison of biometric techniques. Technical report, California Welfare Fraud Prevention System, 2002.

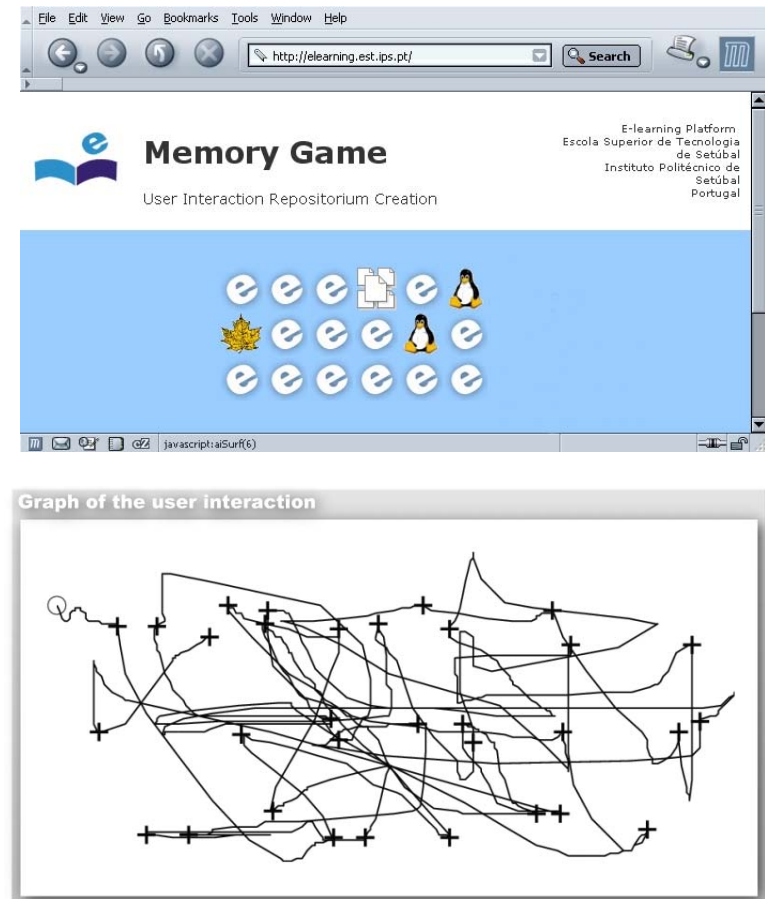# Prevention or Identification of Web Intrusion via Human Computer Interaction Behaviour - A Proposal

Hugo Gamboa, Ana Fred and António Alves Vieira
SCI-158-18

- **Our Proposal**
- **Definitions**
- **The Authentication System**
  - **The Acquisition System**
  - **The Recognition System**
    - **Feature Extraction**
    - **Feature Selection**
    - **Parametric Learning**
    - **Sequential Classification**
- **Results**
- **Conclusion**

Summary

Detect or prevent the intrusion in web sites via the human computer interaction behaviour.

We developed a prototype system accessed via a web browser that records the user interaction while the user is playing the memory game. The system classifies the user as genuine or intruder.

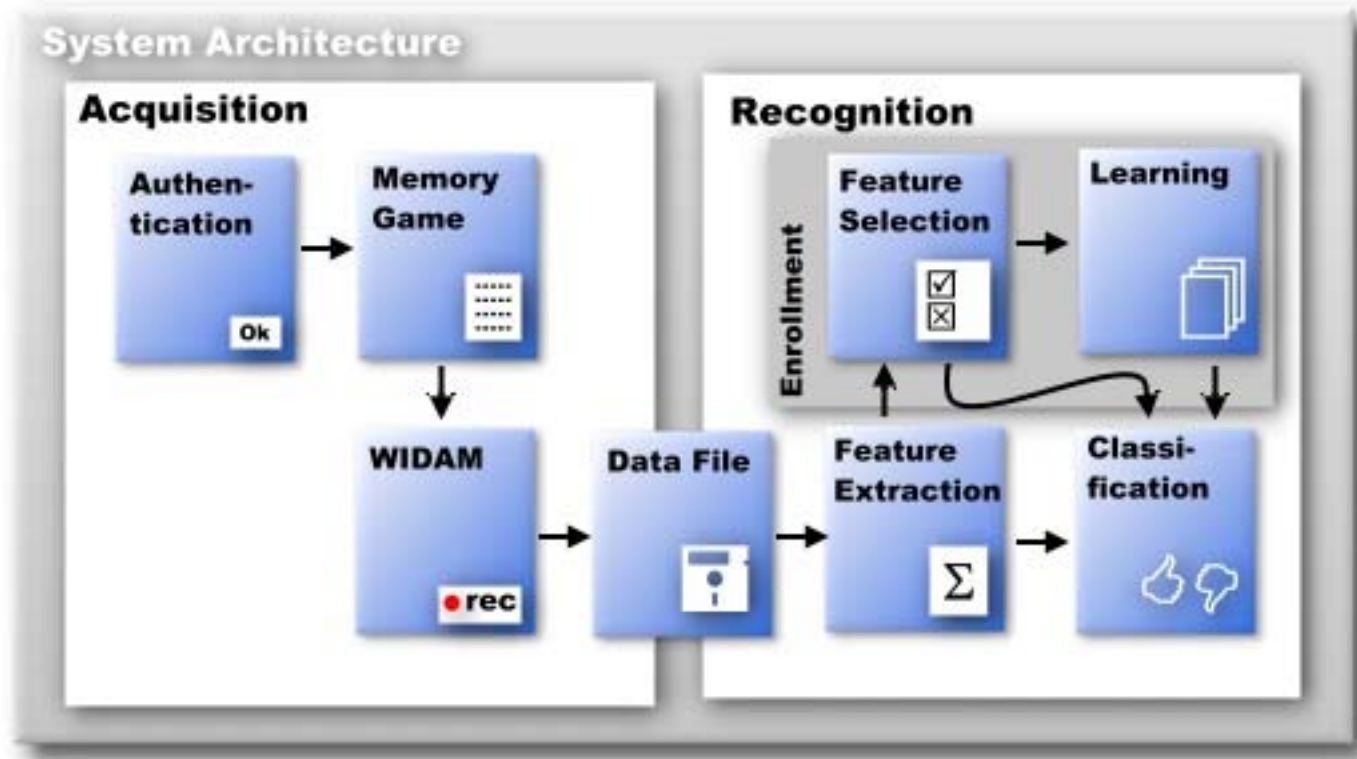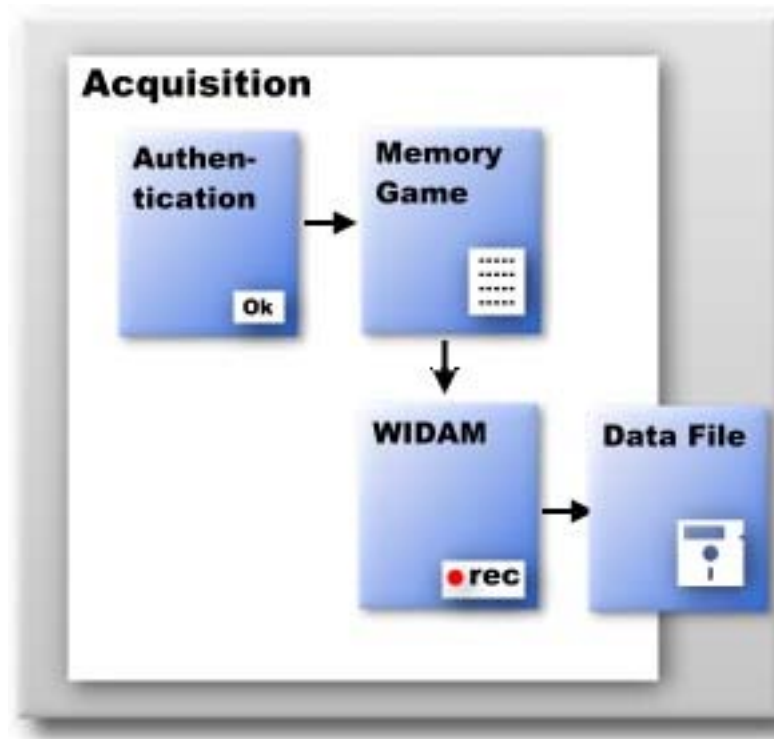A New behavioural biometric technique based on human computer interaction via a pointing device.

File Edit View Go Bookmarks Tools Window Help

http://elearning.est.ips.pt/   Search

Memory Game

E-learning Platform
Escola Superior de Tecnologia
de Setúbal
Instituto Politécnico de
Setúbal
Portugal

User Interaction Repositorium Creation

javascript:aiSurf(6)

Graph of the user interaction

Our Proposal

- **Verification vs Identification**

- **Physical vs Behavioural Biometrics**

- **FRR**
  - **Rate of accesses where a legitim user is rejected by the system.**
- **FAR**
  - **Rate of accesses where an impostor is accepted by the system.**
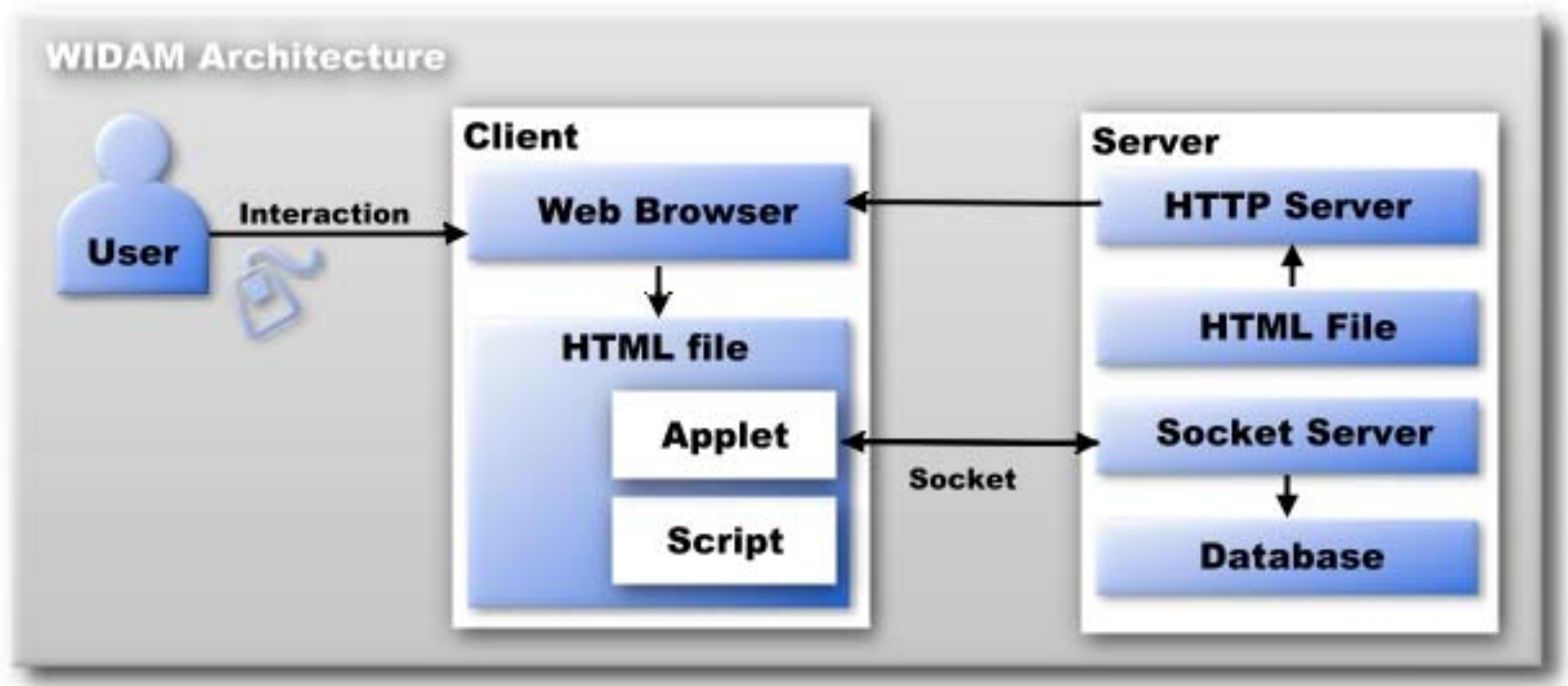- **ERR**
  - **The value at which FAR and FRR are equal**

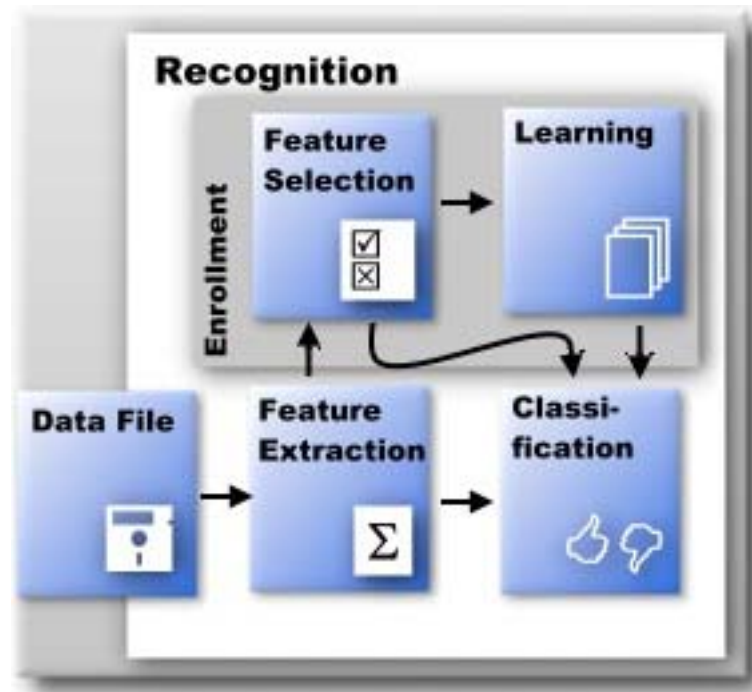
Definitions

The Authentication System

Acquisition System

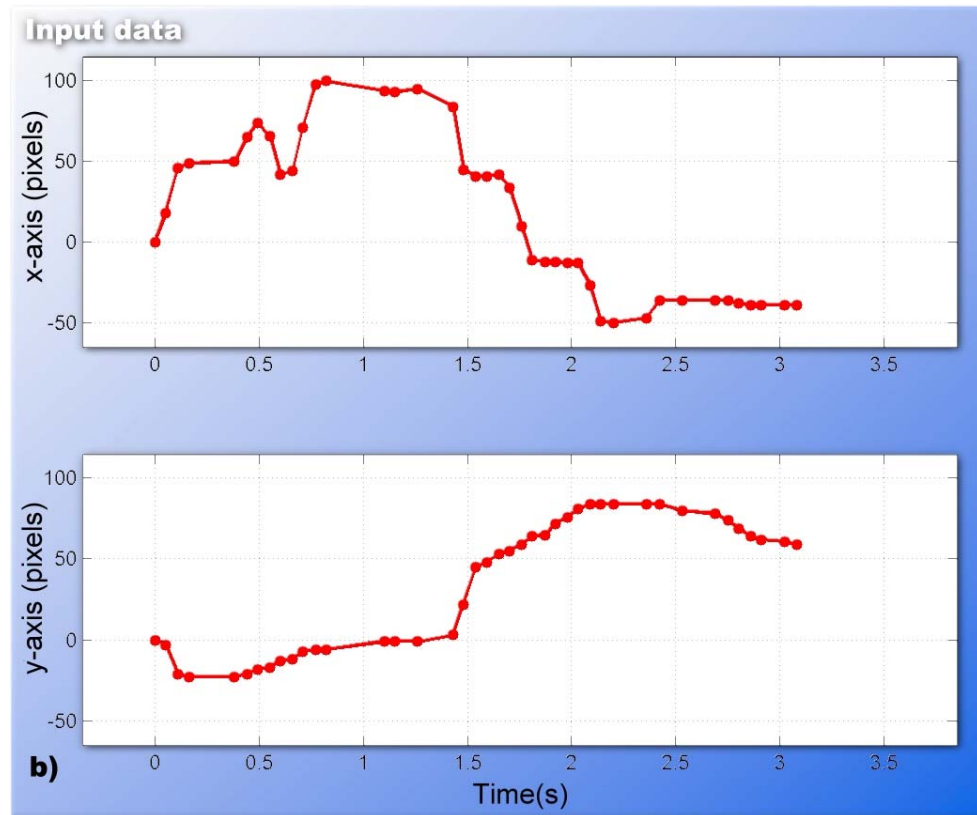# WIDAM – Web Interaction Display and Monitoring

WIDAM Architecture

User → Interaction → Client

Client
- Web Browser
- HTML file
  - Applet
  - Script

Server
- HTTP Server
- HTML File
- Socket Server
- Database

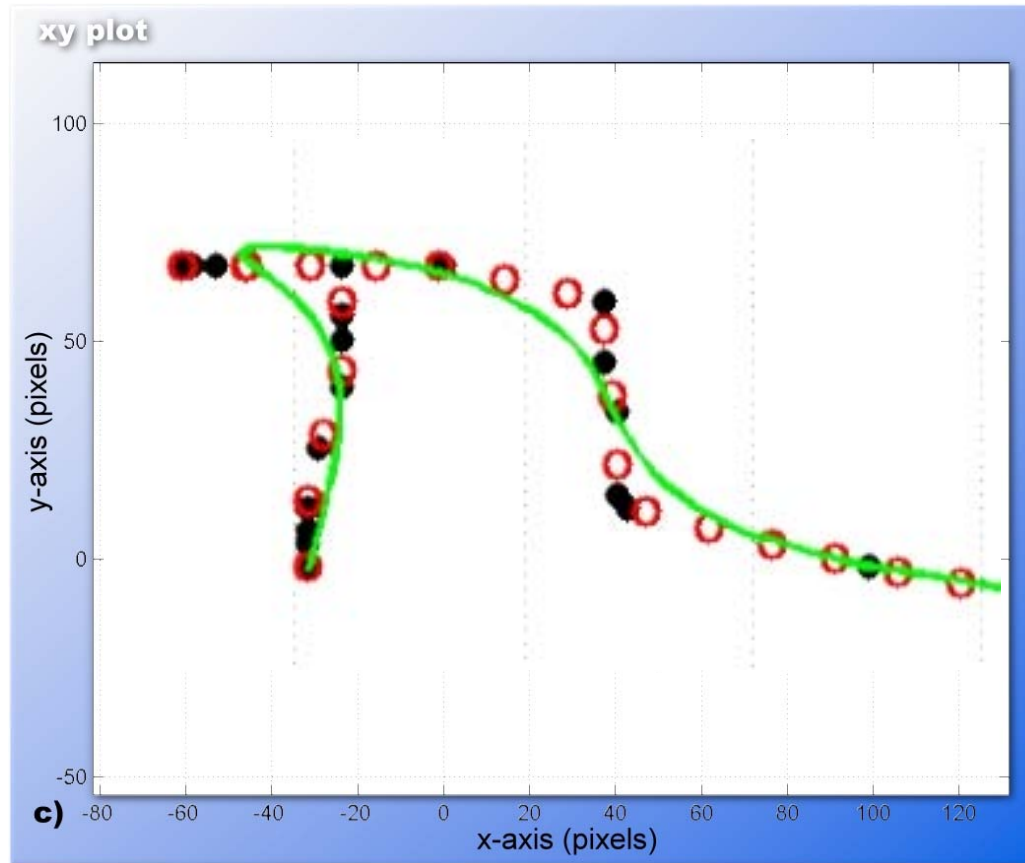Socket

Acquisition System

The Recognition System

A stroke:

Feature Extraction

• Cleaning the input Signal
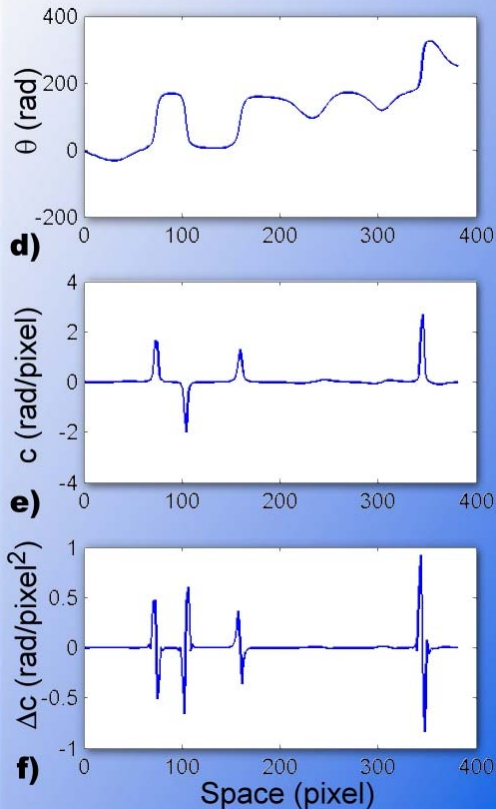

Feature Extraction

Spatial information In the spatial domain we will have six vectors:

- x' — uniformly spaced smoothed horizontal coordinates.

- y' — uniformly spaced smoothed vertical coordinates.

- s' — uniformly spaced path distance from origin.

- θ — angle of the path tangent with the x-axis.

- c — curvature

- Δc — derivative of curvature in order to space.
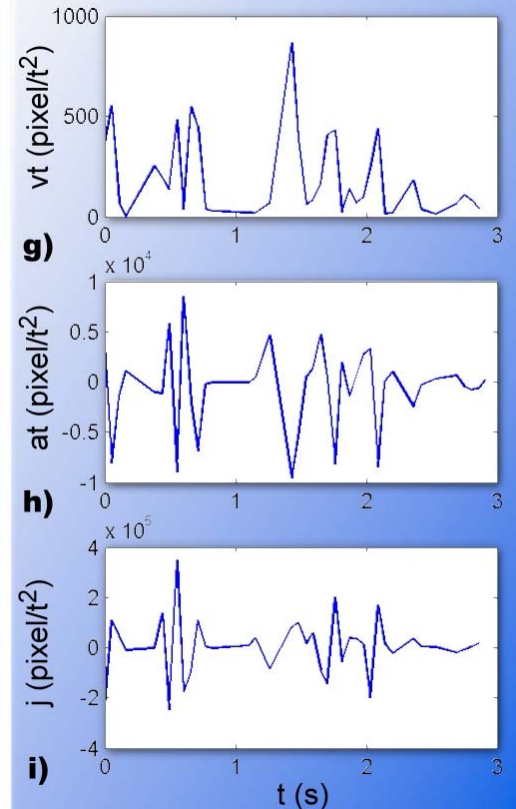
Temporal information In the temporal domain we defined 9 vectors:

- x — the vector with the input $x_i \ldots x_n$ values.

- y — the vector with the input $y_i \ldots y_n$ values.

- t — the input time vector $t_i \ldots t_n$.

- v_x — horizontal velocity.

- v_y — vertical velocity.

- v — tangential velocity.

- $\dot{v}$ — tangential acceleration.

- $\ddot{v}$ — tangential jerk.
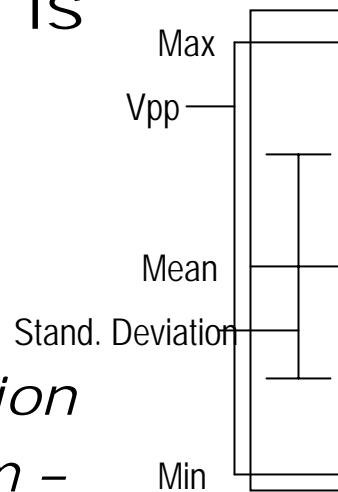
- w — angular velocity.

Feature Extraction

- **From each spatial and temporal vector is extracted:**
  - *minimum*
  - *maximum*
  - *mean*
  - *standard deviation*
  - *range (maximum – minimum)*

Max

Vpp

Mean

Stand. Deviation

Min

- **Other general features:**
  - *Straightness*
  - *Jitter*
  - *Number of high curvature points*
  - *Time to click*
  - *Number of pauses*
  - *Paused time*
  - *Paused time ratio*

A 62-dimensional feature vector is extracted

- *Considering that exists a classifier that receives a subset of features and returns the equal error rate of the system.*

- **Sequential Forward Selection (SFS):**
  - **Selects the best single feature**
  - **Adds one feature at a time to the vector of previously selected features.**
  - **The algorithm stops when the equal error rate does not decrease**

- **User independent feature vector**

1. Create an empty feature subset f_{subset}.
2. Initialize the best equal error rate of the previous interaction, $EER_{last} = 1$.
3. For each feature $f_i,\ \ i = 1...n_{features}$:
   (a) Create the vector with the features to test, $f_{test} = f_{subset} \cup f_i$.
   (b) Set the feature equal error rate (f_{EER_i}) equal to the result of the recognition system test, using the subset f_{test}. $f_{EER_i} = \text{TEST}(f_{test})$.
4. If $\min_i f_{EER_i} > EER_{last}$ exit and return f_{subset}.
5. Set $EER_{last} = \min_i f_{EER_i}$.
6. Set the best feature $f_{best} = \arg\min_i f_{EER_i}$.
7. Set $f_{subset} = f_{test} \cup f_{best}$.
8. Go to 3.

Feature Selection

- **The classification assumes a statistical model for the feature vectors.**
  - **Assuming statistical independence between features:**

  $$p(X \mid user) = \prod p(x_i \mid user)$$

  - **We use the Weibull distribution as a parametric model after a data transformation:**

  $$p_{weibull}(x \mid a, b) = abx^{(b-1)}e^{(-ax^b)}$$

  - **Maximum likelihood estimates of the parameters a and b are obtained**

Parametric Learning
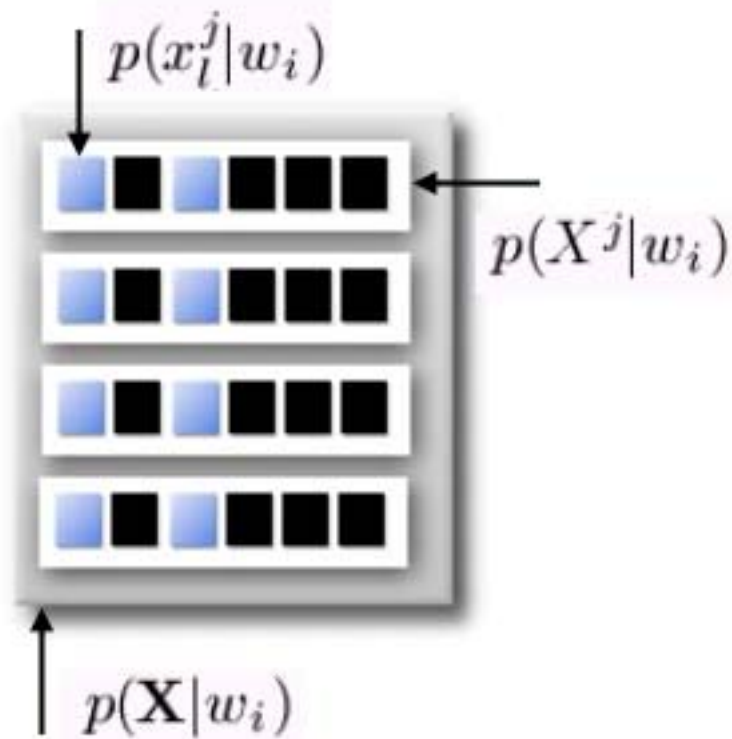
The set of features from user stroke

The set of selected features from the user stroke

$$p(x_l^j|w_i)$$

$$p(X^j|w_i)$$

$$p(\mathbf{X}|w_i)$$

Sequential Classification

Class w_i, $i = 1, \ldots, L$ represent the ith user (L number of users) .
$\mathbf{X} = X^1 \ldots X^{n_s}$, is a sequence of n_s consecutive strokes feature vectors
$X^j = x_1^j \ldots x_{n_{f_i}}^j$ is the feature vector representing the jth stroke,
n_{f_i} is the number of features identified for user w_i

$p(X^j|w_i) = \prod_{l=1}^{n_f} p(x_l^j|w_i)$ (assuming statistical independent features) and
$p(\mathbf{X}|w_i) = \prod_{j=1}^{n_s} p(X^j|w_i)$ (considering stroke independence)

We generate the following two distributions:
$p(\mathbf{X}|w_i)$ (genuine distribution) and $p(\mathbf{X}|\overline{w_i})$ (impostor distribution)

Posterior probability function is

$$p(w_i|\mathbf{X}) = \frac{p(\mathbf{X}|w_i)}{\sum_{k=1}^{L} p(\mathbf{X}|w_k)} = 1 - p(\overline{w_i}|\mathbf{X})$$

Since $p(w_i|\mathbf{X})$ represents an estimate of the probability of the classification
being correct, we establish a *limit*, λ, to select one of the decisions

$$Accept(\mathbf{X} \in w_i) = \begin{cases} true & \text{if } p(w_i|\mathbf{X}) > \lambda \\ false & \text{otherwise} \end{cases}$$
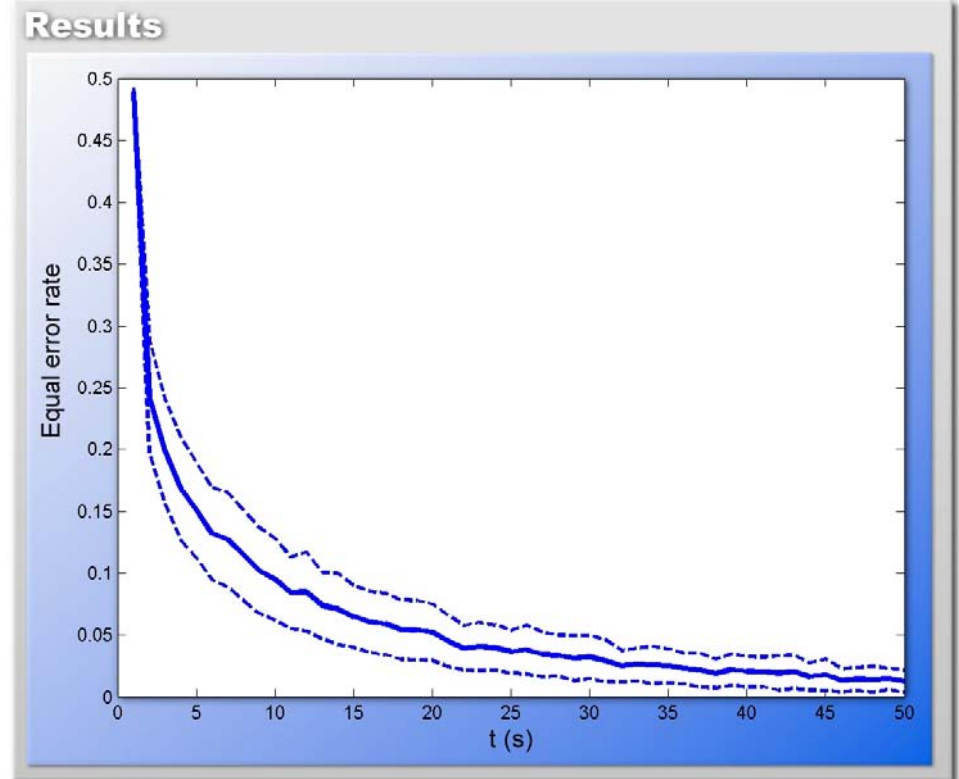
We adjust λ to operate in the equal error rate point to present results.

Sequential Classification

- **Test setup**
  - 50 users
  - 10-15 minutes per user
  - 400 strokes per user

  - Two sub-sets:
    - Training set
      - used in feature selection
      - and parametric learning
    - Testing set
      - used for eer results tests

- **The test**
  - Test the system
    - from 1 to 100 sequential strokes
  - Bootstrap estimates
    - 10000 bootstrap samples from testing set
  - The test returns the EER for each user for each stroke number.

l	eer	std
1	0.480	0.016
2	0.263	0.057
5	0.179	0.056
10	0.118	0.043
20	0.063	0.029
50	0.020	0.016
100	0.007	0.011
200	0.002	0.005

Results

Biometric technic	Equal error rate
Retinal Scan	1:10 000 000
Iris Scan	1:131 000
Fingerprints	1:500
Hand Geometry	1:500
Signature Dynamics	1:50
Voice Dynamics	1:50
$30s$ of User Interaction	1:50
$60s$ of User Interaction	1:100
$90s$ of User Interaction	1:200

Comparison

- **Preliminary results show:**

  - **Comparable performance to other behavioural biometric techniques.**

  - **Inexpensive web based biometric system supported on common available techonologies.**

  - **Cappable of prevention of attacks by webbots by the lack of HCI.**

Conclusion